错误纠正：Python在机器学习中的应用

注意：由于在印刷时出现了很多库在导入时，导致了首字母被大写，所以在运行程序时需要注意，如：sklearn、matplotlib等库，手字母被大写了。

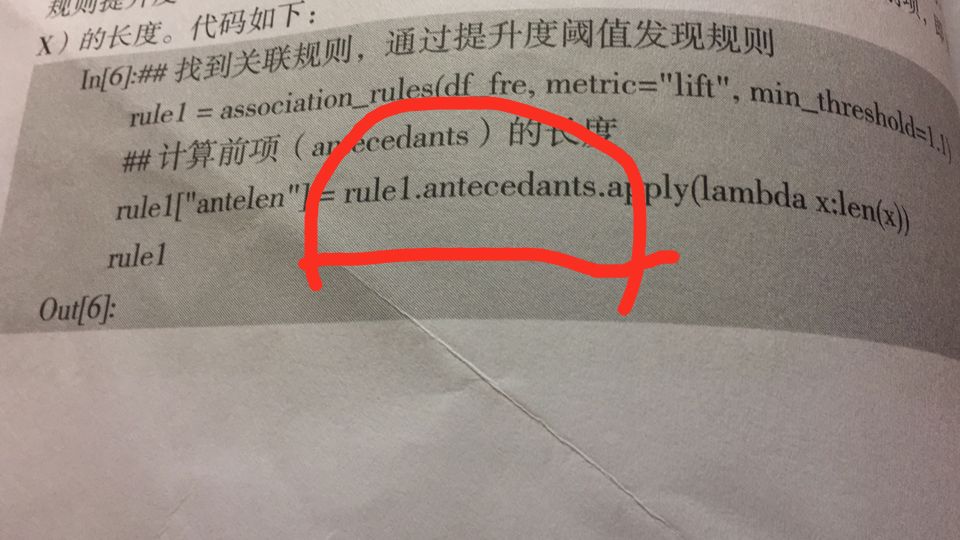
在Python导入库时，一般情况下库的名称应该都是小写字母。

纠错：

1、P77第8行行首缺少了“pos=“。正确的应为：

pos = nx.circular\_layout(G)

2、第六章140页最后一行代码，由于mxltend包版本的更新，较新的版本的输出中添加了更多的内容，并且将原始的输出中antecedants列的列名更换为了antecedents，所以在使用该行程序是，如果mxltend包版本较新，可讲将relu1. antecedants更改为relu1. antecedents。



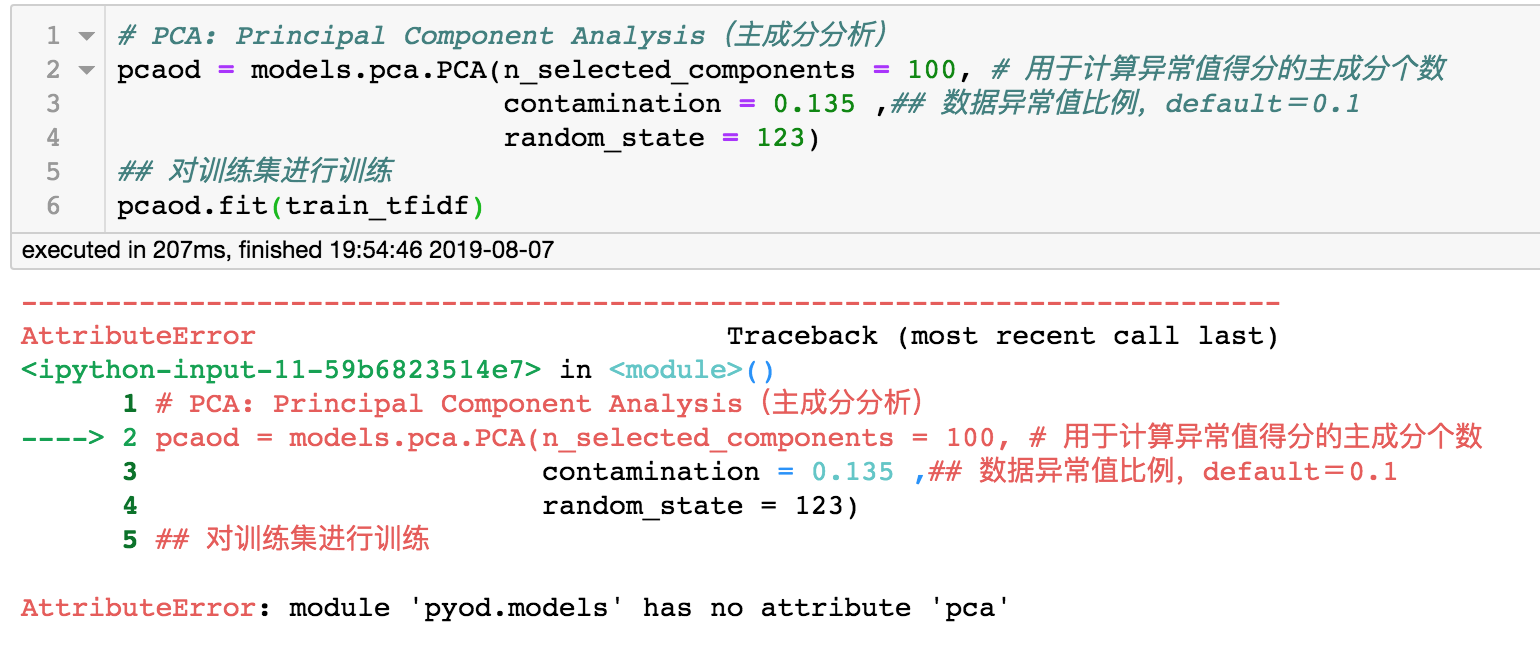
3、第十章 in15模块的程序，原使用的pyod版本为0.5.3

pyod 0.5.3 <pip>

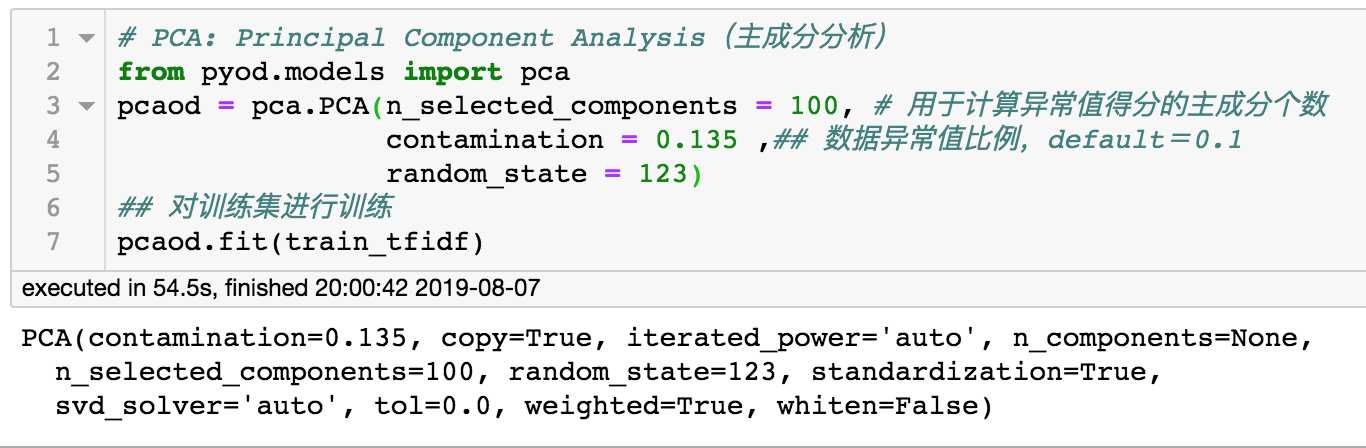
随着时间的发展，现pyod意境更新到0.7.4版本后

pyod 0.7.4 <pip>

运行下图的程序会出错：



使用程序应该更该为：



# PCA: Principal Component Analysis（主成分分析）

from pyod.models import pca

pcaod = pca.PCA(n\_selected\_components = 100, # 用于计算异常值得分的主成分个数

contamination = 0.135 ,## 数据异常值比例，default＝0.1

random\_state = 123)

## 对训练集进行训练

pcaod.fit(train\_tfidf)

同样后面使用孤立森林的算法可使用下面的程序：

## IsolationForest异常值检测器

from pyod.models import iforest

ifood = iforest.IForest(n\_estimators = 5000, ## The number of base estimators in the ensemble.

max\_samples = 1000,

contamination = 0.135 ,## 数据异常值比例，default＝0.1

max\_features = 500,

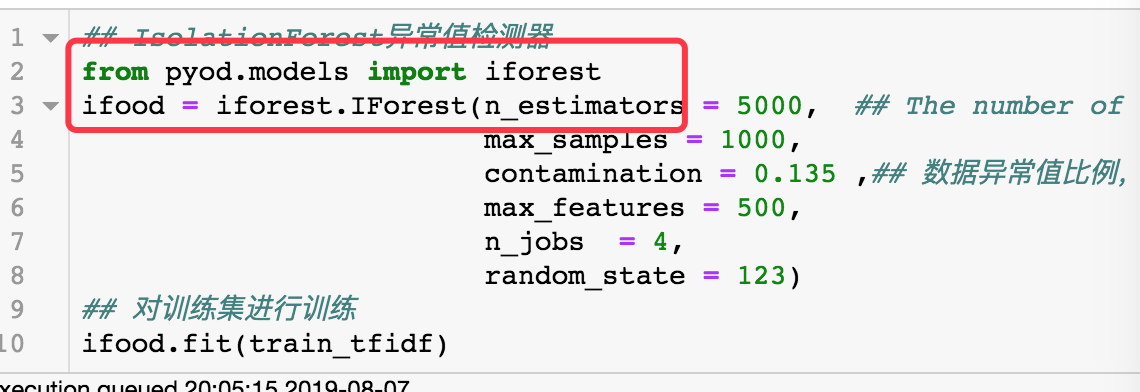
n\_jobs = 4,

random\_state = 123)

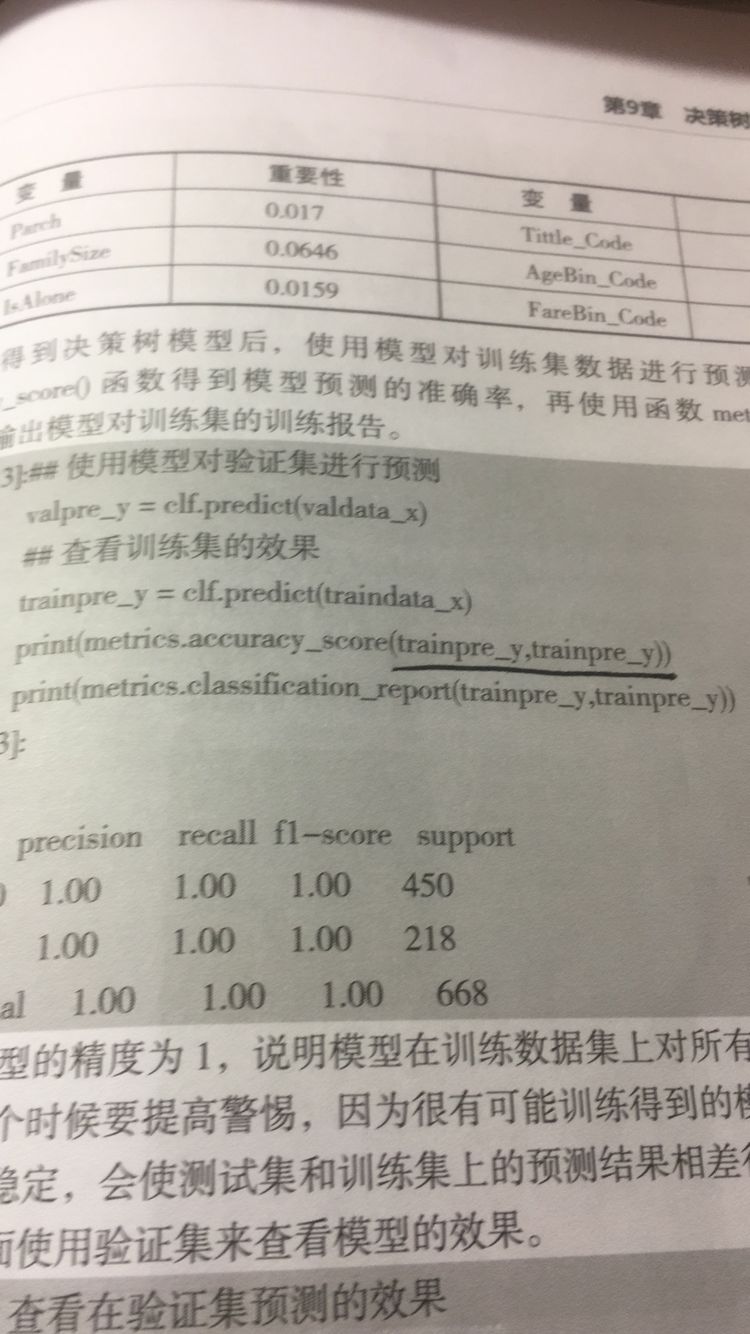
## 对训练集进行训练

ifood.fit(train\_tfidf)

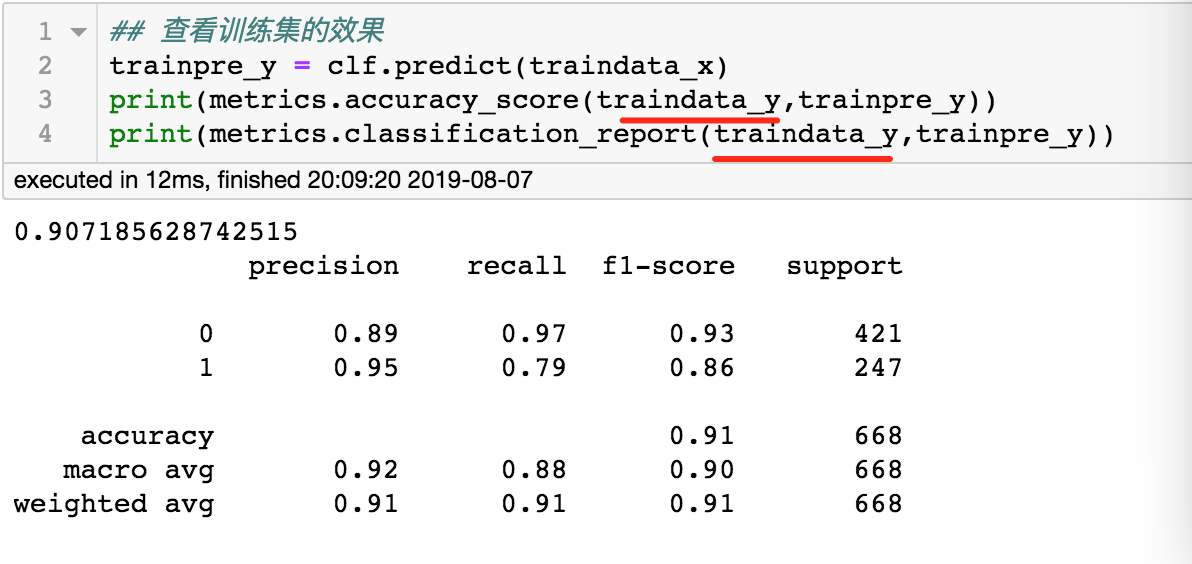
即修改为：



4、第九章 使用决策树对训练数据集进行预测时，计算精度的第一个参数适用错误，错误如下图：



图像中的错误应该修改为如下图所示：并且修改后的结果也如下图所示：

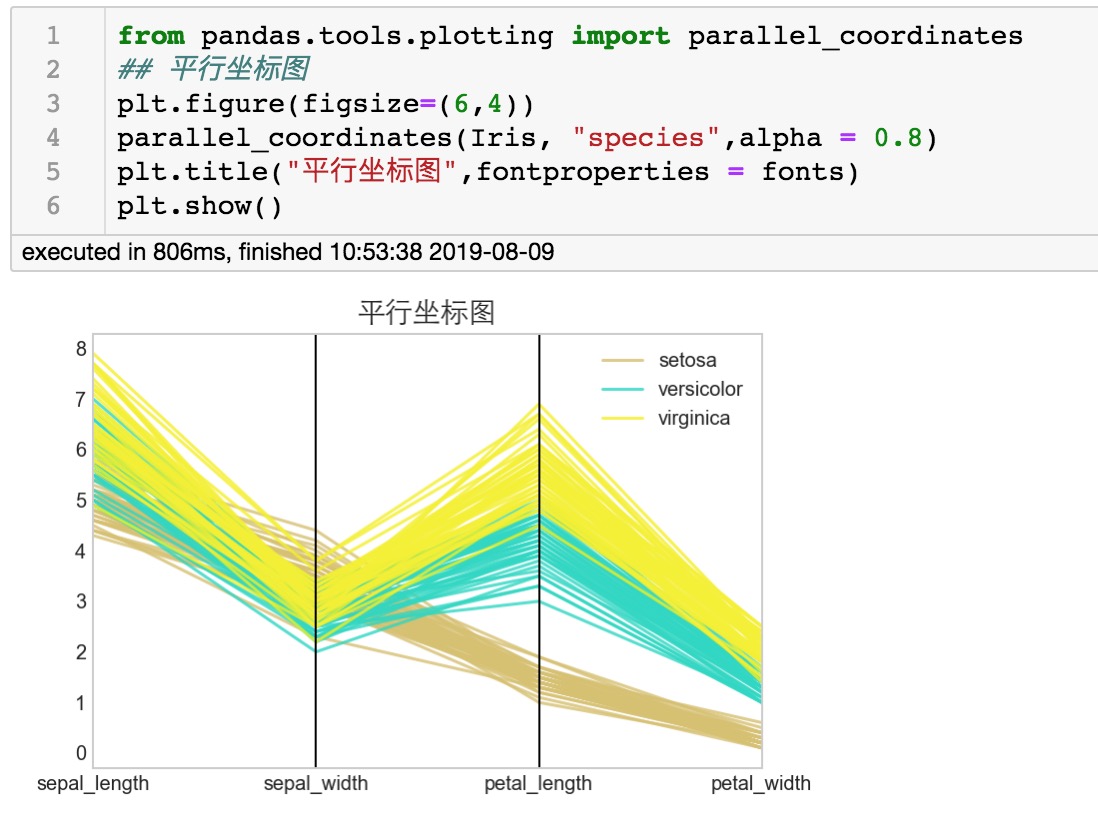


5、第三章 程序片段In28，关于有些pandas的版本无法正确导入的问题，

在使用下面版本：

pandas 0.19.2 <pip>

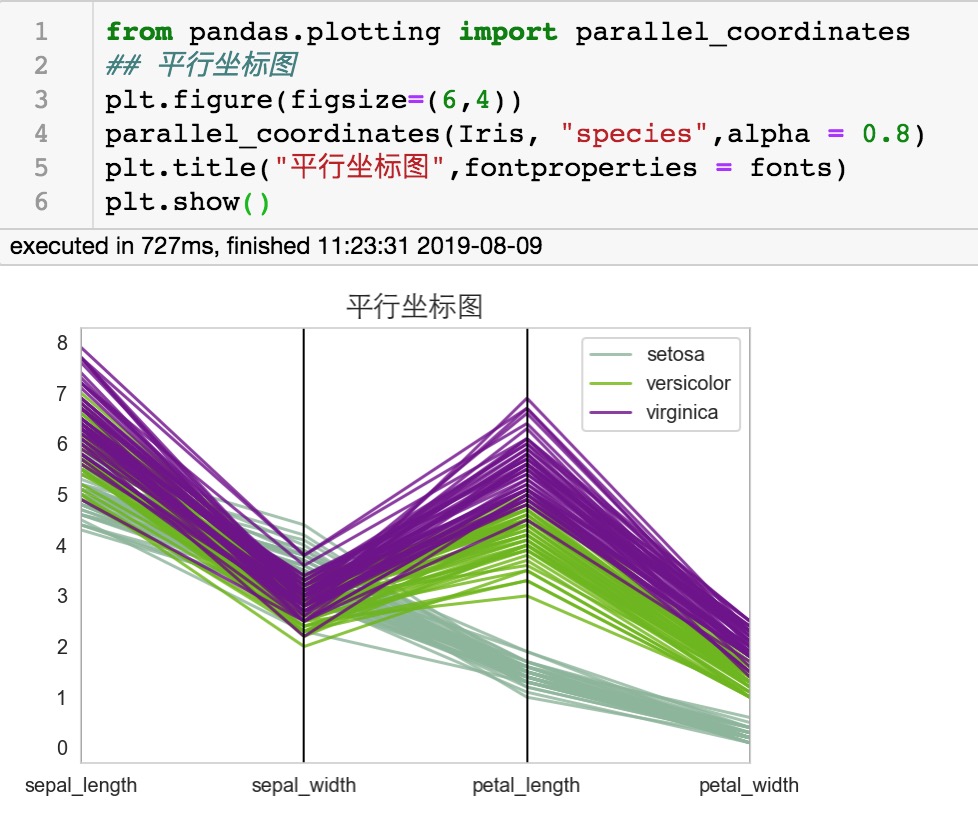
可以正确运行：



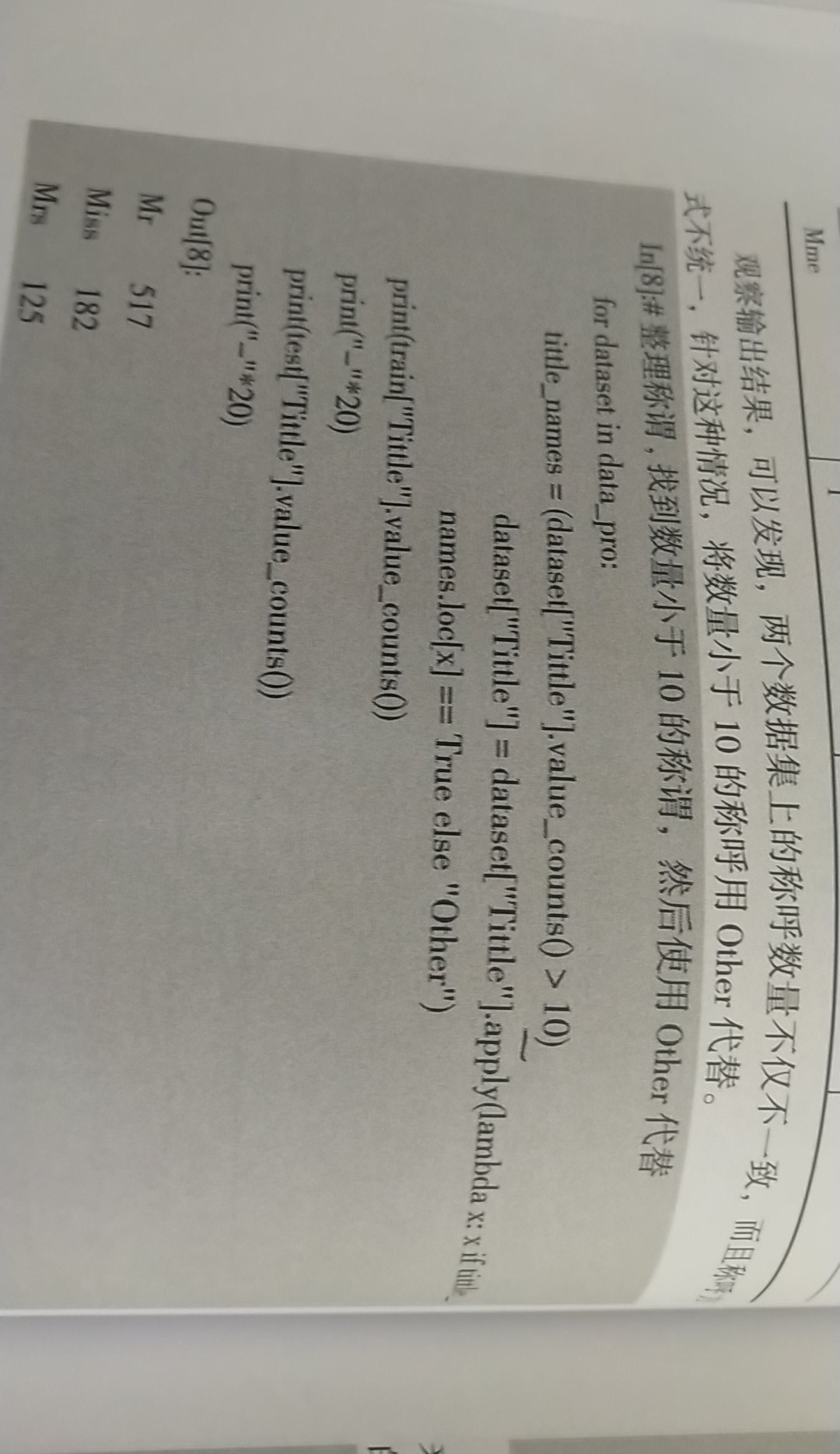
在最新的pandas版本测试时：

pandas 0.25.0 py36h0a44026\_0 anaconda

因为在pandas 0.24.0 之后的版本，需要使用pandas.plotting替换pandas.tools.plotting，替换后的程序及结果如下：



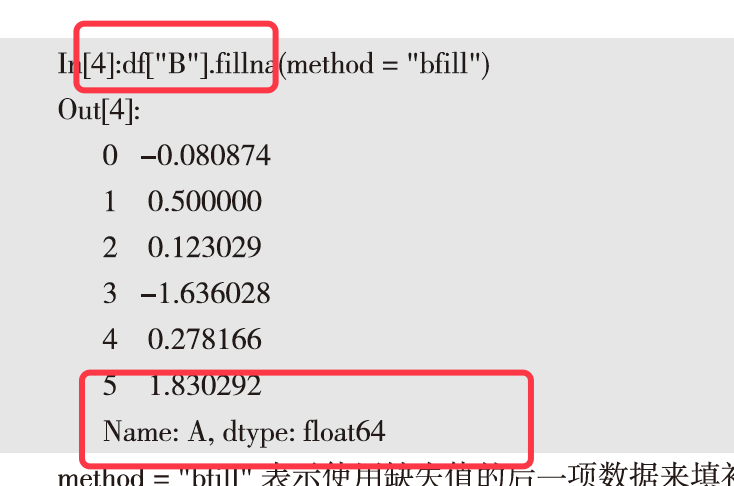
6:第九章程序片段8印刷出现错误，缩进没有调整正确



正确程序片段应为如下图格式所示：

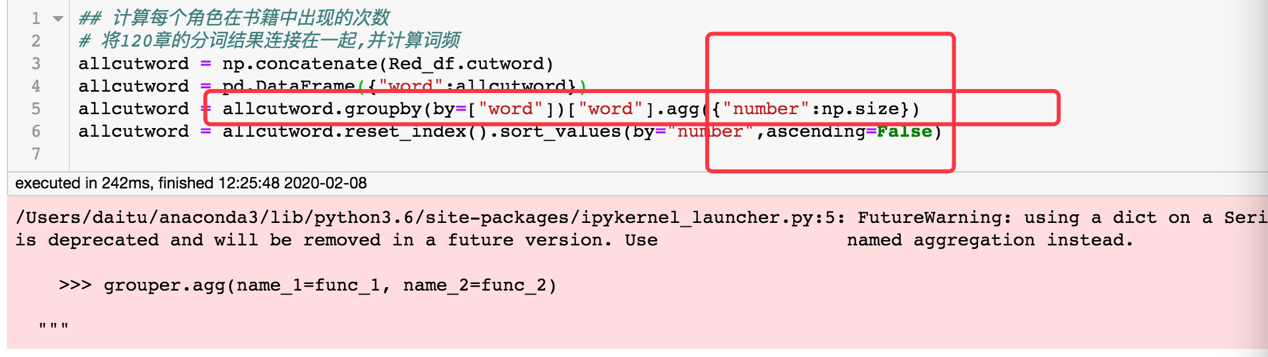


7. 第55页的程序片段4对应的输出结果粘贴错误，需要读者注意一下。



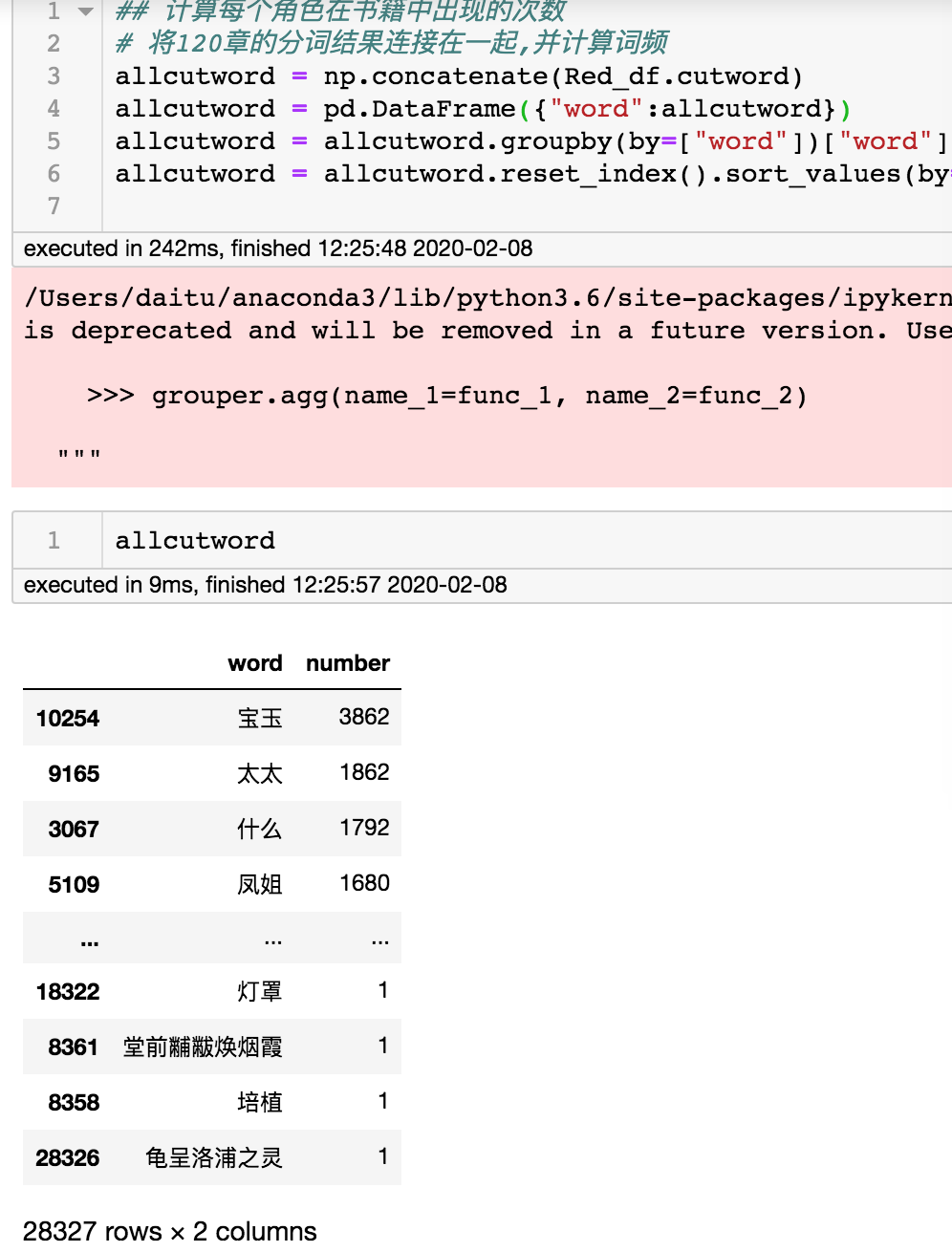
8 书中第185页代码时，遇到这句代码“allcutword = allcutword.groupby(by=["word"])['word'].agg({"number":np.size})”运行时可能会出错，或者出现提醒。

出现原因：在较新的Pandas库中，在使用groupby的agg方法时，意境改变了相关参数的使用方式。例如，运行程序片段后：

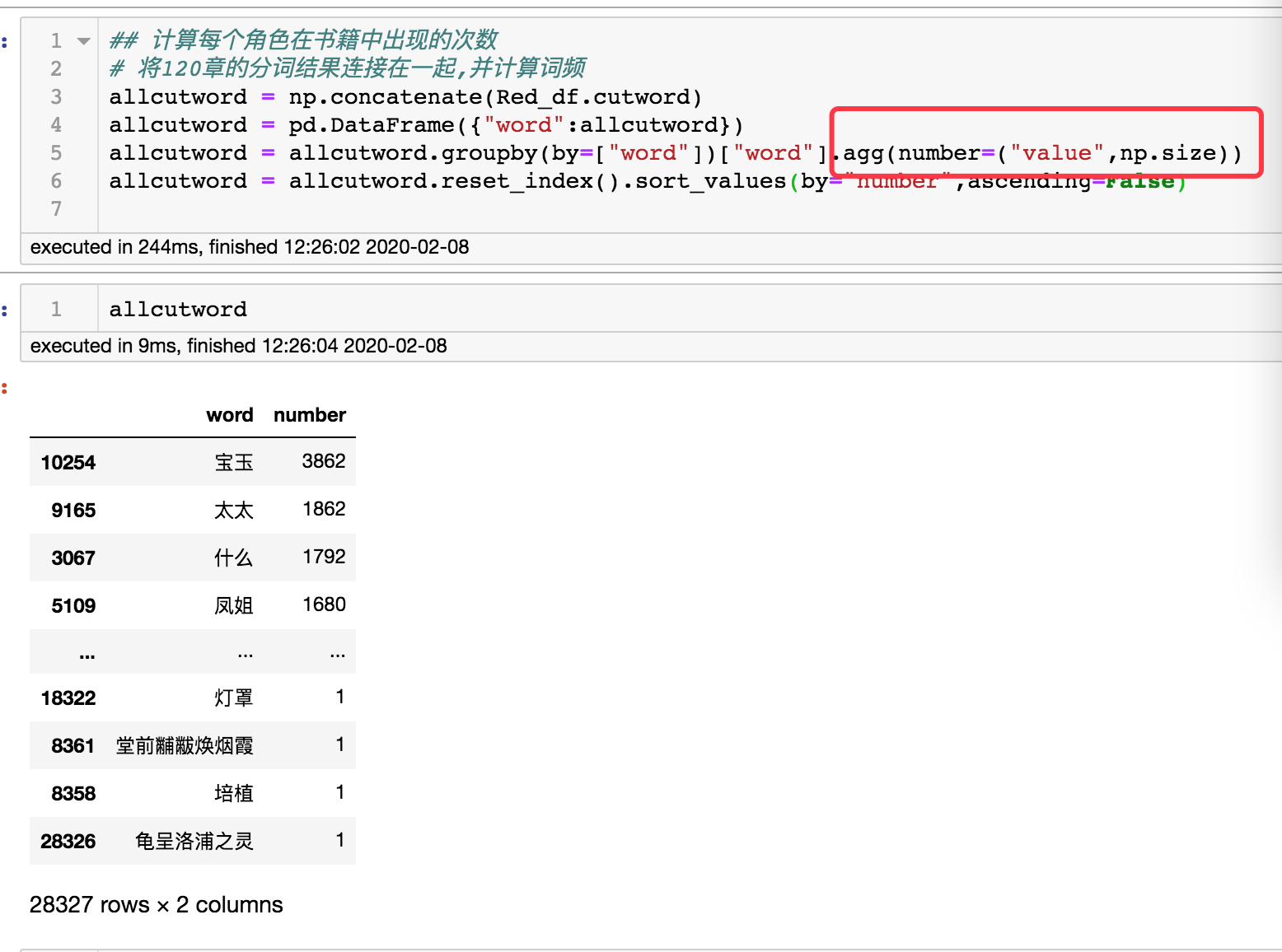


可以注意到，会出现提示相关的用法后期会更新（如果版本Pandas较新，可能会出错）。

我们查看这段程序的运行结果：



错误的修改方式如下：



结果也是正确的。

注意：书中可能会有其它地方也是这样的用法，可能会出现错误，所以相关的程序，按照此方法修改即可。